

KB #7



PCT09

RAW SEQUENCE LISTING

DATE: 06/21/2002

PATENT APPLICATION: US/09/719,017A

TIME: 11:26:37

Input Set : A:\550054.txt

Output Set: N:\CRF3\06212002\I719017A.raw

ENTERED

3 <110> APPLICANT: Pierrard, Jerome
 4 Guitton, Carole
 5 Favre-Bulle, Olivier
 7 <120> TITLE OF INVENTION: Industrial Method For Producing Heterologous Proteins
 8 In E. coli And Strains Useful For Said Method
 10 <130> FILE REFERENCE: 5500*54
 12 <140> CURRENT APPLICATION NUMBER: US 09/719,017A
 C--> 13 <141> CURRENT FILING DATE: 2001-10-09
 15 <150> PRIOR APPLICATION NUMBER: PCT/FR99/01343
 16 <151> PRIOR FILING DATE: 1999-06-08
 18 <150> PRIOR APPLICATION NUMBER: FR 98/07,474
 19 <151> PRIOR FILING DATE: 1998-06-10
 21 <160> NUMBER OF SEQ ID NOS: 6
 23 <170> SOFTWARE: PatentIn Ver. 2.1
 25 <210> SEQ ID NO: 1
 26 <211> LENGTH: 121
 27 <212> TYPE: DNA
 28 <213> ORGANISM: Escherichia coli
 30 <400> SEQUENCE: 1
 31 gaattccctg ttgacaatta atcatcgaac tagttaacta gtacgcagct tggctgcagg 60
 32 tcgacctgca gccaaagcttg ggcatacatt caatcaattg ttatctaagg aaatacttac 120
 33 a 121
 36 <210> SEQ ID NO: 2
 37 <211> LENGTH: 1793
 38 <212> TYPE: DNA
 39 <213> ORGANISM: Alcaligenes faecalis
 41 <220> FEATURE:
 42 <221> NAME/KEY: CDS
 43 <222> LOCATION: (123)..(1190)
 45 <400> SEQUENCE: 2
 46 gaattccctg ttgacaatta atcatcgaac tagttaacta gtacgcagct tggctgcagg 60
 48 tcgacctgca gccaaagcttg ggcatacatt caatcaattg ttatctaagg aaatacttac 120
 50 at atg cag aca aga aaa atc gtc cgg gca gcc gcc gta cag gcc gcc 167
 51 Met Gln Thr Arg Lys Ile Val Arg Ala Ala Val Gln Ala Ala
 52 1 5 10 15
 54 tct ccc aac tac gat ctg gca acg ggt gtt gat aaa acc att gag ctg 215
 55 Ser Pro Asn Tyr Asp Leu Ala Thr Gly Val Asp Lys Thr Ile Glu Leu
 56 20 25 30
 58 gct cgt cag gcc cgc gat gag ggc tgt gac ctg atc gtg ttt ggt gaa 263
 59 Ala Arg Gln Ala Arg Asp Glu Gly Cys Asp Leu Ile Val Phe Gly Glu
 60 35 40 45
 62 acc tgg ctg ccc ggc tat ccc ttc cac gtc tgg ctg ggc gca ccg gcc 311
 63 Thr Trp Leu Pro Gly Tyr Pro Phe His Val Trp Leu Gly Ala Pro Ala

RAW SEQUENCE LISTING

DATE: 06/21/2002

PATENT APPLICATION: US/09/719,017A

TIME: 11:26:37

Input Set : A:\550054.txt

Output Set: N:\CRF3\06212002\I719017A.raw

64	50	55	60	
66	tgg tcg ctg aaa tac agt gcc cgc tac tat gcc aac tcg ctc tcg ctg	359		
67	Trp Ser Leu Lys Tyr Ser Ala Arg Tyr Tyr Ala Asn Ser Leu Ser Leu			
68	65 70 75			
70	gac agt gca gag ttt caa cgc att gcc cag gcc gca cgg acc ttg ggt	407		
71	Asp Ser Ala Glu Phe Gln Arg Ile Ala Gln Ala Ala Arg Thr Leu Gly			
72	80 85 90 95			
74	att ttc atc gca ctg ggt tat agc gag cgc agc ggc ggc agc ctt tac	455		
75	Ile Phe Ile Ala Leu Gly Tyr Ser Glu Arg Ser Gly Gly Ser Leu Tyr			
76	100 105 110			
78	ctg ggc caa tgc ctg atc gac gac aag ggc cag atg ctg tgg tcg cgt	503		
79	Leu Gly Gln Cys Leu Ile Asp Asp Lys Gly Gln Met Leu Trp Ser Arg			
80	115 120 125			
82	cgc aaa ctc aaa cct aca cat gtt gag cgc acc gtg ttt ggt gaa ggt	551		
83	Arg Lys Leu Lys Pro Thr His Val Glu Arg Thr Val Phe Gly Glu Gly			
84	130 135 140			
86	tat gcc cga gat ctg att gtg tcc gac acc gag ctg ggc cgc gtc ggt	599		
87	Tyr Ala Arg Asp Leu Ile Val Ser Asp Thr Glu Leu Gly Arg Val Gly			
88	145 150 155			
90	gcc ctg tgc tgc tgg gag cac ctg tcc ccc ttg agc aag tac gcg ctg	647		
91	Ala Leu Cys Cys Trp Glu His Leu Ser Pro Leu Ser Lys Tyr Ala Leu			
92	160 165 170 175			
94	tac tcc cag cac gaa gcc att cac att gcc gcc tgg ccg tcc ttt tcg	695		
95	Tyr Ser Gln His Glu Ala Ile His Ile Ala Ala Trp Pro Ser Phe Ser			
96	180 185 190			
98	ctg tac agc gaa cag gcc cat gcg ctc agc gcc aag gtg aac atg gct	743		
99	Leu Tyr Ser Glu Gln Ala His Ala Leu Ser Ala Lys Val Asn Met Ala			
100	195 200 205			
102	gcc tcg caa atc tat tcg gtt gaa ggc cag tgc ttt acc atc gcc gcc	791		
103	Ala Ser Gln Ile Tyr Ser Val Glu Gly Gln Cys Phe Thr Ile Ala Ala			
104	210 215 220			
106	agc agt gtc gtc acc cag gag aca ctg gac atg ctg gaa gta ggt gaa	839		
107	Ser Ser Val Val Thr Gln Glu Thr Leu Asp Met Leu Glu Val Gly Glu			
108	225 230 235			
110	cac aac gcc tcc ctg ctg aaa gtg ggc ggc ggc agt tcc atg att ttt	887		
111	His Asn Ala Ser Leu Leu Lys Val Gly Gly Gly Ser Ser Met Ile Phe			
112	240 245 250 255			
114	gcg ccg gac gga cgc aca ttg gct ccc tac ctg cca cac gat gcc gaa	935		
115	Ala Pro Asp Gly Arg Thr Leu Ala Pro Tyr Leu Pro His Asp Ala Glu			
116	260 265 270			
118	ggc ctg atc att gcc gat ctg aac atg gaa gaa att gcc ttc gcc aag	983		
119	Gly Leu Ile Ile Ala Asp Leu Asn Met Glu Glu Ile Ala Phe Ala Lys			
120	275 280 285			
122	gcg atc aac gac cct gtg ggc cac tac tcc aaa ccc gag gcc acc cgt	1031		
123	Ala Ile Asn Asp Pro Val Gly His Tyr Ser Lys Pro Glu Ala Thr Arg			
124	290 295 300			
126	ctg gta ctg gac ctg ggg cac cgt gag ccc atg act cgg gtg cat tcc	1079		
127	Leu Val Leu Asp Leu Gly His Arg Glu Pro Met Thr Arg Val His Ser			
128	305 310 315			

RAW SEQUENCE LISTING

DATE: 06/21/2002

PATENT APPLICATION: US/09/719,017A

TIME: 11:26:37

Input Set : A:\550054.txt

Output Set: N:\CRF3\06212002\I719017A.raw

```

130 aaa agc gtg atc cag gaa gaa gct ccc gag ccg cac gtg caa agt acg 1127
131 Lys Ser Val Ile Gln Glu Glu Ala Pro Glu Pro His Val Gln Ser Thr
132 320 325 330 335
134 gct gcg ccc gtc gcc gtc agc cag act cag gac tcg gat acg cta ctg 1175
135 Ala Ala Pro Val Ala Val Ser Gln Thr Gln Asp Ser Asp Thr Leu Leu
136 340 345 350
138 gtg caa gaa ccg tcc tgacccccaaa agatgacaag gcccgggcaa actgtccggg 1230
139 Val Gln Glu Pro Ser
140 355
142 tcttgattcc ttctgcgtcc cggatccact agttctagag tcgacctgca ggcattgcaag 1290
144 cttgggtccc acctgacccc atgccgaact cagaagtgaac acgccgtagc gccgatggta 1350
146 gtgtgggggtc tccccatgcg agagtaggga actgccaggc atcaaataaa acgaaaggct 1410
148 cagtcgaaaag actgggcctt tcgttttatac tgttgttgt cggtgaacgc tctcctgagt 1470
150 aggacaaatc cgccgggagc ggatttgaac gttgcgaagc aacggccccg aggggtggcg 1530
152 gcaggacgcc cgccataaac tgccaggcat caaattaagc agaaggccat cctgacggat 1590
154 ggcctttttg cgttttctaca aactcttctc gtcgtcatat ctacaagcca tccccccaca 1650
156 gatacggtaa actagcctcg tttttgcatac aggaaagcag ctatgaacca ctcttataaa 1710
158 ccctggaaca catttggcac tgatcataat gctcagcaca ttgtatgtgc cgaagacgaa 1770
160 caacaattac tcaatgcccg cgg 1793
163 <210> SEQ ID NO: 3
164 <211> LENGTH: 356
165 <212> TYPE: PRT
166 <213> ORGANISM: Alcaligenes faecalis
168 <400> SEQUENCE: 3
169 Met Gln Thr Arg Lys Ile Val Arg Ala Ala Val Gln Ala Ala Ser
170 1 5 10 15
172 Pro Asn Tyr Asp Leu Ala Thr Gly Val Asp Lys Thr Ile Glu Leu Ala
173 20 25 30
175 Arg Gln Ala Arg Asp Glu Gly Cys Asp Leu Ile Val Phe Gly Glu Thr
176 35 40 45
178 Trp Leu Pro Gly Tyr Pro Phe His Val Trp Leu Gly Ala Pro Ala Trp
179 50 55 60
181 Ser Leu Lys Tyr Ser Ala Arg Tyr Tyr Ala Asn Ser Leu Ser Leu Asp
182 65 70 75 80
184 Ser Ala Glu Phe Gln Arg Ile Ala Gln Ala Ala Arg Thr Leu Gly Ile
185 85 90 95
187 Phe Ile Ala Leu Gly Tyr Ser Glu Arg Ser Gly Gly Ser Leu Tyr Leu
188 100 105 110
190 Gly Gln Cys Leu Ile Asp Asp Lys Gly Gln Met Leu Trp Ser Arg Arg
191 115 120 125
193 Lys Leu Lys Pro Thr His Val Glu Arg Thr Val Phe Gly Glu Gly Tyr
194 130 135 140
196 Ala Arg Asp Leu Ile Val Ser Asp Thr Glu Leu Gly Arg Val Gly Ala
197 145 150 155 160
199 Leu Cys Cys Trp Glu His Leu Ser Pro Leu Ser Lys Tyr Ala Leu Tyr
200 165 170 175
202 Ser Gln His Glu Ala Ile His Ile Ala Ala Trp Pro Ser Phe Ser Leu
203 180 185 190
205 Tyr Ser Glu Gln Ala His Ala Leu Ser Ala Lys Val Asn Met Ala Ala

```

RAW SEQUENCE LISTING

DATE: 06/21/2002

PATENT APPLICATION: US/09/719,017A

TIME: 11:26:37

Input Set : A:\550054.txt

Output Set: N:\CRF3\06212002\I719017A.raw

```

206          195          200          205
208 Ser Gln Ile Tyr Ser Val Glu Gly Gln Cys Phe Thr Ile Ala Ala Ser
209          210          215          220
211 Ser Val Val Thr Gln Glu Thr Leu Asp Met Leu Glu Val Gly Glu His
212 225          230          235          240
214 Asn Ala Ser Leu Leu Lys Val Gly Gly Gly Ser Ser Met Ile Phe Ala
215          245          250          255
217 Pro Asp Gly Arg Thr Leu Ala Pro Tyr Leu Pro His Asp Ala Glu Gly
218          260          265          270
220 Leu Ile Ile Ala Asp Leu Asn Met Glu Glu Ile Ala Phe Ala Lys Ala
221          275          280          285
223 Ile Asn Asp Pro Val Gly His Tyr Ser Lys Pro Glu Ala Thr Arg Leu
224          290          295          300
226 Val Leu Asp Leu Gly His Arg Glu Pro Met Thr Arg Val His Ser Lys
227 305          310          315          320
229 Ser Val Ile Gln Glu Glu Ala Pro Glu Pro His Val Gln Ser Thr Ala
230          325          330          335
232 Ala Pro Val Ala Val Ser Gln Thr Gln Asp Ser Asp Thr Leu Leu Val
233          340          345          350
235 Gln Glu Pro Ser
236          355
240 <210> SEQ ID NO: 4
241 <211> LENGTH: 35
242 <212> TYPE: DNA
243 <213> ORGANISM: Artificial Sequence
245 <220> FEATURE:
246 <223> OTHER INFORMATION: Description of Artificial Sequence:Primer
248 <400> SEQUENCE: 4
249 ccccccgcaca gctgtcttgc ttctgaattt ctgcc 35
252 <210> SEQ ID NO: 5
253 <211> LENGTH: 26
254 <212> TYPE: DNA
255 <213> ORGANISM: Artificial Sequence
257 <220> FEATURE:
258 <223> OTHER INFORMATION: Description of Artificial Sequence:Primer
260 <400> SEQUENCE: 5
261 ttgacgtcag tagctgaaca ggaggg 26
264 <210> SEQ ID NO: 6
265 <211> LENGTH: 12
266 <212> TYPE: DNA
267 <213> ORGANISM: Artificial Sequence
269 <220> FEATURE:
270 <223> OTHER INFORMATION: Description of Artificial Sequence:linker
272 <400> SEQUENCE: 6
273 aatacttaca cc 12

```

VERIFICATION SUMMARY

DATE: 06/21/2002

PATENT APPLICATION: US/09/719,017A

TIME: 11:26:38

Input Set : A:\550054.txt

Output Set: N:\CRF3\06212002\I719017A.raw

L:13 M:271 C: Current Filing Date differs, Replaced Current Filing Date